Figure 1:

```
PREF-1      MIATGALLRVLLLLLAFGHS---TYGAECDPPCDPQYGFCEADNVCRCHVGWEGPLCDKC
SCS0009     MPSGCRCLHLVCLLCILGAPGQPVRADDCSSHCDLAHGCCAPDGSCRCDPGWEGLHCERC
            * .   *... **  *        .*  **  * *  *. ***  ****  *..*

PREF-1      VTAPGCVNGVCKEPWQCICKDGWDGKFCEIDVRACTSTPCANNGTCVDLEKGQYECSCTP
SCS0009     VRMPGCQHGTCHQPWQCICHSGWAGKFCDK------------------------------
            *  *** .* *..******. ** ****.
```

Additional exon coding for 30 aa is inserted in
SECP59 (INCYTE WO0226982) HKAAV61 (HGS WO0134768)
PRO40 (GENENTECH WO0224888) PREF-1 (HYSEQ WO0157233)
LP230S (LILLY WO0248361) AAG67516 (SKB WO0166690)
35.31 (BIODOOR WO0226809)

```
PREF-1      GFSGKDCQHKAGPCV INGSPCQHGGACVDDEGQASHASCLCPPGFSGNFCEIVAATNSCT
SCS0009     GFHGRDCERKAGPCEQAGSPCRNGGQCQDDQGFALNFTCRCLVGFVGARCEVN --VDDCL
            ** *.**..*****   ****..** * **.* * . .* *  ** *  **.      *

PREF-1      PNPCENDGVCTDIGGDFRCRCPAGFVDKTCSRPVSNCAS GPCQNGGTCLQHTQVSFECLC
SCS0009     MRPCANGATCLDGINRFSCLCPEGFAGRFCTINLDDCASRPCQRGARCRDRVHD -FDCLC
            ** *   * *  . * * ** **  . *.  .  *** *** *  *  . .  *.***
```

Initial Met in NOV8 (CURAGEN WO0255704)

```
PREF-1      KPPFMGPTCAKKRGAS-PVQVTHLPSGYGLTYRLTPGVHELPVQQPEQHILKVSMKELN-
SCS0009     PSGYGGKTCELVLPVPDPPTTVDTPLGP-TSAVVVPATGPAPHSAG-AGLLRISVKEVVR
               . * **       *     * *   . . *     *       .*..*.**.

PREF-1      KSTPLLTEGQAICFTILGVLTSLVVLGTVAIVFLNKCETWV SNLRYNHTFRKKKNLLLQY
SCS0009     RQEAGLGEPSLVALVVFGALTAALVLATVLLTLR ----AWRRGVCPPGPCCYPAPHYAPA
            .    * *   ..  . * **. .** ** .       .*  ..

PREF-1      NSGEELAVNIIFPEKIDMTTFNKEAGDEEI-
SCS0009     CQDQECQVSMLPAGLPL PRDLPPEPGKTTAL
               .*  * ..          * *
```

Figure 2:

Clustal W alignment of predicted amino acid sequences of SCS0009 prediction and splice variants SV3, SV4 and SV5.

```
CLUSTAL W (1.83) multiple sequence alignment

SV4             MP SGCRCLHLVCLLCILGAPGQPVRADDCSSHCDLAHGCCAPDGSCRCDPGWEGLHCERC
SV5             MPSGCRCLHLVCLLCILGAPGQPVRADDCSSHCDLAHGCCAPDGSCRCDPGWEGLHCERC
SV3             MPSGCRCLHLVCLLCILGAPGQPVRADDCSSHCDLAHGCCAPDGSCRCDPGWEGLHCERC
SCS0009         MPSGCRCLHLVCLLCILGAPGQPVRAD DCSSHCDLAHGCCAPDGSCRCDPGWEGLHCERC
                ************************************************************

SV4             VRMPGCQHGTCHQPWQCICHSGWAGKFCDKDEHICTTQSPCQNGGQCMYDGGGEYHCVCL
SV5             VRMPGCQHGTCHQPWQCICHSGWAGKFCDKDEHICTTQSPCQNGGQCMYDG GGEYHCVCL
SV3             VRMPGCQHGTCHQPWQCICHSGWA ------DEHICTTQSPCQNGGQCMYDGGGEYHCVCL
SCS0009         VRMPGCQHGTCHQPWQCICHSGWAGKFCDK ----------------------------

SV4              PGFHGRDCERKAGP ----------------------------------------------
SV5              PGFHGRDCERKAGPCEQAGSPCRNGGQCQDDQGFALNFTCRCLVGFVGARCEVNVDDCLM
SV3              PGFHGRDCERKAGPCEQAGSPCRNGGQCQDDQGFALNFTCRCLVGFVGARCEVNVDDCLM
SCS0009          -GFHGRDCERKAGPCEQAGSPCRNGGQCQDDQGFALNFTCRCLVGFVGARCEVNVDDCLM

SV4              ---------------------- ---------------------------------------
SV5             RPCANGATCLDGINRFSCLCPEGFAGRFCTINLDDCASRPCQRGARCRDRVHDFDCLCPS
SV3             RPCANGATCLDGINRFSCLCPEGFAGRFCTINLDDCASRPCQRGARCRDRVHDFDCLCPS
SCS0009         RPCANGATCLDGINRFSCLCPEGFAGRFCTINLDDCASRPCQRGARC RDRVHDFDCLCPS

SV4              ------------------------------------------------------------
SV5             GYGGKTCELVLPVPDPPTTVDTPLGPTSAVVVPATGPAPHSAGAGLLRISVKEVVRRQEA
SV3             GYGGKTCELVLPVPDPPTTVDTPLGPTSAVVVPATGPAPHSAGAGLLRISVKEVVRRQEA
SCS0009         GYGGKTCELVLPVPDPPTTVDTPLGPTSAVVVPATGPAPHSAGAGLLRISVKEVVRRQEA
```

Figure 3:

Clustal W alignment of precited amino acid sequences of SCS0009 with SV1 and SV2

```
SCS0009          ---------------------------------- MPSGCRCLHLVCLLCILGAPG QPVRA
SV1-ORF          ------------------------------------------------------------
SV2-ORF          ------------------------------------------------------------

SCS0009          DDCSSHCDLAHGCCAPDGSC RCDPGWEGLHCERCVRMPGCQHGTCHQPWQCICHSGWAGK
SV1-ORF          ------------------------ ---------- MPGCQHGTCHQPWQCICHSGWAGK
SV2-ORF          ---------------------------------- MPGCQHGTCHQPWQCICHSGWA --
                                                     **********************

SCS0009          FCDK------------------------------- GFHGRDCERKAG PCEQAGSPCRNGG
SV1-ORF          FCDKDEHICTTQSPCQNGGQCMYDGGGEYHCVCLPGFHGRDCERKAGPCEQAGSPCRNGG
SV2-ORF          -----DEHICTTQSPCQNGGQCMYDGGGEYHCVCLPGFHGRDCERKAGPCEQAGSPCRNGG
                     ************************************************************

SCS0009          QCQDDQGFALNFTCRCLVGFVGARCEVNVDDCLMRPCANGATCLDGINRFSCLCPEGFAG
SV1-ORF          QCQDDQGFALNFTCRCLVGFVGARCEVNVDDCLMRPCANGATCLDGINRFSCLCPEGFAG
SV2-ORF          QCQDDQGFALNFTCRCLVGFVGARCEVNVDDCLMRPCANGATCLDGINRFSCLCPEGFAG
                 ****************** ****************************************

SCS0009          RFCTINLDDCASRPCQRGARCRDRVHDFDCLCPSGYGGKTCELVLPVPDPPTTVDTPLGP
SV1-ORF          RFCTINLDDCASRPCQRGARCRDRVHDFDCLCPSGYGGKTCELVLPVPDPPTTVDTPLGP
SV2-ORF          RFCTINLDDCASRPCQRGARCRDRVHDFDCLCPSGYGGKTCE LVLPVPDPPTTVDTPLGP
                 ************************************************************

SCS0009          TSAVVVPATGPAPHSAGAGLLRISVKEVVRRQEAGLGEPSLVALVVFGALTAALVLATVL
SV1-ORF          TSAVVVPATGPAPHSAGAGLLRISVKEVVRRQEAGLGEPSLVALVVFGALTAALVLATVL
SV2-ORF          TSAVVVPATGPAPHSAGAGLLRISVKEVVRRQEAGLGEPSLVALVVFGALTAALVLATVL
                 ************************************************************

SCS0009          LTLRAWRRGVCPPGPCCYPAPHYAPACQDQECQVSMLPAGLPLPRDLPPEPGKTTAL .
SV1-ORF          LTLRAWRRGVCPPGP CCYPAPHYAPACQDQECQVSMLPAGLPLPRDLPPEPGKTTAL .
SV2-ORF          LTLRAWRRGVCPPGPCCYPAPHYAPACQDQECQVSMLPAGLPLPRDLPPEPGKTTAL .
                 ************************************************************
```

XY = exon boundaries.

In the translation, the SV1 and SV2 se quences are shown representing the longest ORF available.

The predicted signal peptide of SCS0009 is shown highlighted in yellow.
The SV1 and SV2 longest ORFs do not contain predicted signal peptides.

Figure 4:

Nucleotide sequence of SCS0009 prediction with translation

```
   1  AGACGGCAAC GTGGACAGGA AGAAGCGGAG GGCGAGGAGG AGCAGAGGAG CACACAGATG
  61  AAGCAGGTGT CCACGCGTCC GGCCGTCCAT CCGTCCGTCC CTCCTGGGGC CGGCGCTGAC
 121  CATGCCCAGC GGCTGCCGCT GCCTGCATCT CGTGTGCCTG TTGTGCATTC TGGGGGCTCC
        M  P  S  G  C  R  C  L  H  L  V  C  L  L  C  I  L  G  A

 181  CGGTCAGCCT GTCCGAGCCG ATGACTGCAG CTCCCACTGT GACCTGGCCC ACGGCTGCTG
        P  G  Q  P  V  R  A  D  D  C  S  S  H  C  D  L  A  H  G  C

 241  TGCACCTGAC GGCTCCTGCA GGTGTGACCC GGGCTGGGAG GGG CTGCACT GTGAGCGCTG
        C  A  P  D  G  S  C  R  C  D  P  G  W  E  G  L  H  C  E  R

 301  TGTGAGGATG CCTGGCTGCC AGCACGGTAC CTGCCACCAG CCATGGCAGT GCATCTGCCA
        C  V  R  M  P  G  C  Q  H  G  T  C  H  Q  P  W  Q  C  I  C

 361  CAGTGGCTGG GC AGGCAAGT TCTGTGACAA AGGCTTCCAT GGGCGTGACT GCGAGCGCAA
        H  S  G  W  A  G  K  F  C  D  K  G  F  H  G  R  D  C  E  R

 421  GGCTGGACCC TGTGAACAGG CAGGCTCCCC ATGCCGCAAT GGCGGGCAGT GCCAGGACGA
        K  A  G  P  C  E  Q  A  G  S  P  C  R  N  G  G  Q  C  Q  D

 481  CCAGGGCTTT GCTCTCAACT TCACGTGCCG CTGCTTGGTG GGCTTTGTGG GTGCCCGCTG
        D  Q  G  F  A  L  N  F  T  C  R  C  L  V  G  F  V  G  A  R

 541  TGAGGTAAAT GTGGATGACT GCCTGATGCG GCCTTGTGCT AACGGTGCCA CCTGCCTTGA
        C  E  V  N  V  D  D  C  L  M  R  P  C  A  N  G  A  T  C  L

 601  CGGCATAAAC CGCTTCTCCT GCCTCTGTCC TGAGGGCTTT GCTGGACGCT TCTGCACCAT
        D  G  I  N  R  F  S  C  L  C  P  E  G  F  A  G  R  F  C  T

 661  CAACCTGGAT GACTGTGCCA GCCGCCCATG CCAGAGAGGG GCCCGCTGTC G GGACCGTGT
        I  N  L  D  D  C  A  S  R  P  C  Q  R  G  A  R  C  R  D  R

 721  CCACGACTTC GACTGCCTCT GCCCCAGTGG CTATGGTGGC AAGACCTGTG AGCTTGTCTT
        V  H  D  F  D  C  L  C  P  S  G  Y  G  G  K  T  C  E  L  V

 781  ACCTGTCCCA GACCCCCCAA CCACAGTGGA CACCCCTCTA GGGCCCACCT CAGCTGTAGT
        L  P  V  P  D  P  P  T  T  V  D  T  P  L  G  P  T  S  A  V

 841  GGTACCTGCC ACGGGGCCAG CCCCCCACAG CGCAGGGGCT GGTCTGCTGC GGATCTCAGT
        V  V  P  A  T  G  P  A  P  H  S  A  G  A  G  L  L  R  I  S

 901  GAAGGAGGTG GTGCGGAGGC AAGAGGCTGG GCTAGGTGAG CCTAGCTTGG TGGCCCTGGT
        V  K  E  V  V  R  R  Q  E  A  G  L  G  E  P  S  L  V  A  L

 961  GGTGTTTGGG GCCCTCACTG CTGCCCTGGT TCTGGCTACT GTGTTGCTGA CCCTGAGGGC
        V  V  F  G  A  L  T  A  A  L  V  L  A  T  V  L  L  T  L  R

1021  CTGGCGCCGG GGTGTCTGCC CCCCTGGACC CTGTTGCTAC CCTGCCCCAC ACTATGCTCC
        A  W  R  R  G  V  C  P  P  G  P  C  C  Y  P  A  P  H  Y  A

1081  AGCGTGCCAG GACCAGGAGT GTCAGGTTAG CATGCTGCCA GCAGGGCTCC CCCTGCCACG
        P  A  C  Q  D  Q  E  C  Q  V  S  M  L  P  A  G  L  P  L  P

1141  TGACTTGCCC CCTGAGCCTG GAAAGACCAC AGCACTGTGA TGGAGGTGGG GGCTTTCTGG
        R  D  L  P  P  E  P  G  K  T  T  A  L
```

```
1201  CCCCCTTCCT CACCTCTTCC ACCCCTCAGA CTGGAGTGGT CCGTTCTCA C CACCCTTCAG
1261  CTTGGGTACA CACACAGAGG AGACCTCAGC CTCACACCAG AAATATTATT TTTTTAATAC
1321  ACAGAATGTA AGATGGAATT TTATCAAATA AAACTATGAA AATGCAAGTG GGCTCCTATG
1381  CCAGAAAAAC CCACCTGGCG TTCCAGATGC AAGAGGGCCA GAGCAGAGGC CTGGTTCTGG
1441  GGAAGCCTCA GGATGC TGCC CACCAAGGAG TGATTTCCAA AGAGTAATCC AGGGTGCCCT
1501  TTTCCCTTCT GGGGAAGTGT GGAGAGGTAG AGCCCCAGAG GAGAATGTAA ACAAGCAGCC
1561  AGCACCTCTG TATAGGCCCG GCCTGGATCA GAGAGAGGGG AGAACTCTGC AGGGTGTGGG
1621  ATTGGGCTCA GGGACCTCCG AGTGAGGCAG GGACTCCCTG CTG
```

Figure 5:

Nucleotide sequence with translation of cDNA insert in image clone 5478078 (SCS0009-SV3)

```
   1  ATGCCCAGCG GCTGCCGCTG CCTGCATCTC GTGTGCCTGT TGTGCATTCT GGGGGCTCCC
        M  P  S   G  C  R   C  L  H  L   V  C  L   L  C  I   L  G  A  P

  61  GGTCAGCCTG TCCGAGCCGA TGACTGCAGC TCCCACTGTG ACCTGGCCCA CGGCTGCTGT
        G  Q  P   V  R  A   D  D  C  S   S  H  C   D  L  A   H  G  C  C

 121  GCACCTGACG GCTCCTGCAG GTGTGACCCG GGCTGGGAGG GGCTGCACTG TGAGCGCTGT
        A  P  D   G  S  C   R  C  D  P   G  W  E   G  L  H   C  E  R  C

 181  GTGAGGATGC CTGGCTGCCA GCACGGTACC TGCCACCAGC CATGGCAGTG CATCTGCCAC
        V  R  M   P  G  C   Q  H  G  T   C  H  Q   P  W  Q   C  I  C  H

 241  AGTGGCTGGG CAGATGAACA TATCTGTACC ACGCAGTCCC CCTGCCAGAA TGGAGGCCAG
        S  G  W   A  D  E   H  I  C  T   T  Q  S   P  C  Q   N  G  G  Q

 301  TGCATGTATG ACGGGGGCGG TGAGTACCAT TGTGTGTGCT TACCAGGCTT CCATGGGCGT
        C  M  Y   D  G  G   G  E  Y  H   C  V  C   L  P  G   F  H  G  R

 361  GACTGCGAGC GCAAGGCTGG ACCCTGTGAA  CAGGCAGGCT CCCCATGCCG CAATGGCGGG
        D  C  E   R  K  A   G  P  C  E   Q  A  G   S  P  C   R  N  G  G

 421  CAGTGCCAGG ACGACCAGGG CTTTGCTCTC AACTTCACGT GCCGCTGCTT GGTGGGCTTT
        Q  C  Q   D  D  Q   G  F  A  L   N  F  T   C  R  C   L  V  G  F

 481  GTGGGTGCCC GCTGTGAGGT AAATGTGGAT GACTGCCTGA TGCGGCCTTG TGCTAACGGT
        V  G  A   R  C  E   V  N  V  D   D  C  L   M  R  P   C  A  N  G

 541  GCCACCTGCC TTGACGGCAT AAACCGCTTC TCCTGCCTCT GTCCTGAGGG CTTTGCTGGA
        A  T  C   L  D  G   I  N  R  F   S  C  L   C  P  E   G  F  A  G

 601  CGCTTCTGCA CCATCAACCT GGATGACTGT GCCAGCCGCC CATGCCAGAG AGGGGCCCGC
        R  F  C   T  I  N   L  D  D  C   A  S  R   P  C  Q   R  G  A  R

 661  TGTCGGGACC GTGTCCATGA CTTCGACTGC CTCTGCCCCA GTGGCTATGG TGGCAAGACT
        C  R  D   R  V  H   D  F  D  C   L  C  P   S  G  Y   G  G  K  T

 721  TGTGAGCTTG TCTTACCTGT CCCAGACCCC CCAACCACAG TGGACACCCC TCTAGGGCCC
        C  E  L   V  L  P   V  P  D  P   P  T  T   V  D  T   P  L  G  P

 781  ACCTCAGCTG TAGTGGTACC TGCCAC GGGG CCAGCCCCCC ACAGCGCAGG GGCTGGTCTG
        T  S  A   V  V  V   P  A  T  G   P  A  P   H  S  A   G  A  G  L

 841  CTGCGGATCT CAGTGAAGGA GGTGGTGCGG AGGCAAGAGG CTGGGCTAGG TGAGCCTAGC
        L  R  I   S  V  K   E  V  V  R   R  Q  E   A  G  L   G  E  P   S

 901  TTGGTGGCCC TGGTGGTGTT TGGGGCCCTC ACTGCTGCCC TGGTTCTGGC TACTGTGTTG
        L  V  A   L  V  V   F  G  A  L   T  A  A   L  V  L   A  T  V  L

 961  CTGACCCTGA GGGCCTGGCG CCGGGGTGTC TGCCCCCCTG GACCCTGTTG CTACCCTGCC
        L  T  L   R  A  W   R  R  G  V   C  P  P   G  P  C   C  Y  P  A

1021  CCACACTATG CTCCAGCGTG CCAGGACCAG GAGTGTCAGG TTAGCATGCT GCCAGCAGGG
        P  H  Y   A  P  A   C  Q  D  Q   E  C  Q   V  S  M   L  P  A  G

1081  CTCCCCCTGC CACGTGACTT GCCCCCTGAG CCTGGAAAGA CCACAGCACT G
        L  P  L   P  R  D   L  P  P  E   P  G  K   T  T  A   L
```

Figure 6 :

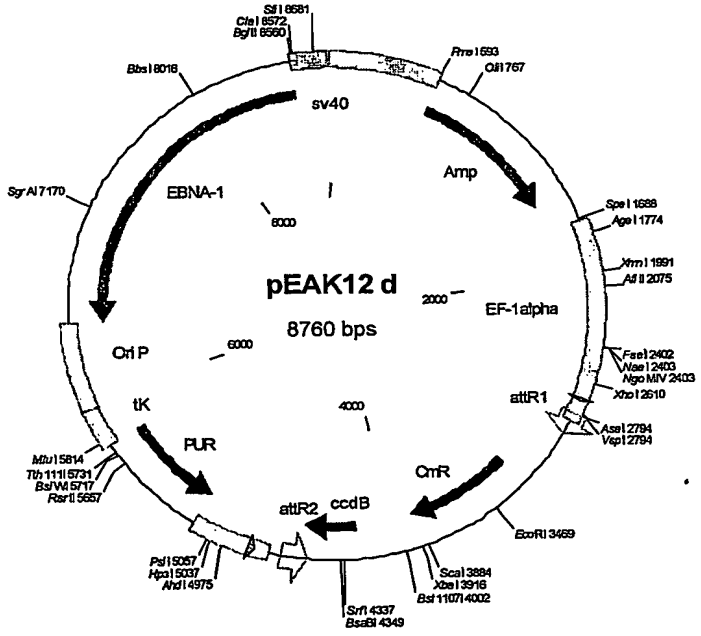
Alignment of predicted amino acid sequence of SCS0009 with SCS0009-SV3

```
SCS0009      MPSGCRCLHLVCLLCILGAPGQPVRADDCSSHCDLAHGCCAPDGSCRCDPGWEGLHCERC
SCS0009SV3   MPSGCRCLHLVC LLCILGAPGQPVRADDCSSHCDLAHGCCAPDGSCRCDPGWEGLHCERC
             ************************************************************

SCS0009      VRMPGCQHGTCHQPWQCICHSGWAGK ------------------------ FCDKGFHGR
SCS0009SV3   VRMPGCQHGTCHQPWQCICHSGWADEHICTTQSPCQ NGGQCMYDGGGEYHCVCLPGFHGR
             ************************.:                          .*  *****

SCS0009      DCERKAGPCEQAGSPCRNGGQCQDDQGFALNFTCRCLVGFVGARCEVNVDDCLMRPCANG
SCS0009SV3   DCERKAGPCEQAGSPCRNGGQCQDDQGFALNFTCRCLVGFVGARCEVNVDDCLMRPCANG
             ************************************************************

SCS0009      ATCLDGINRFSCLCPEGFAGRFCTINLDDCASRPCQRGARCRDRVHDFDCLCPSGYGGKT
SCS0009SV3   ATCLDGINRFSCLCPEGFAGRFCTINLDDCASRPCQRGARCRDRVHDFDCLCPSGYGGKT
             ******* ****************************************************

SCS0009      CELVLPVPDPPTTVDTPLGPTSAVVVPATGPAPHSAGAGLLRISVKEVVRRQEAGLGEPS
SCS0009SV3   CELVLPVPDPPTTVDTPLGPTSAVVVPATGPAPHSAGAGLLRISVKEVVRRQEAGLGEPS
             ******************************** ****************************

SCS0009      LVALVVFGALTAALVLATVLLTLRAWRRGVCPPGPCCYPAPHYAPACQDQECQVSMLPAG
SCS0009SV3   LVALVVFGALTAALVLATVLLTLRAWRRGVCPPGPCCYPAPHYAPACQDQECQVSMLPAG
             ************************************************************ *****

SCS0009      LPLPRDLPPEPGKTTAL
SCS0009SV3   LPLPRDLPPEPGKTTAL
             *****************
```

Figure 7:

Map of expression vector pEAK12d
Molecule: pEAK12 d, 8760 bps DNA Circular
File Name: pEAK12DEST.cm5
Description: Mammali an cell expression vector (plasmid ID 11345)
Molecule Features:

Type	Start	End	Name	Description
REGION	2	595		pmb -ori
GENE	596	1519	Amp	
REGION	1690	2795	EF -1alpha	
REGION	2703	2722		position of pEAK12F primer
REGION	2796	2845		MCS
MARKER	2855		attR1	
GENE	3256	3915	CmR	
GENE	4257	4562	ccdB	
MARKER	4603	C	attR2	
REGION	4733	4733		MCS
REGION	4734	5162		poly A/splice
REGION	4819	4848 C		position of pEAK12R primer
GENE	5781	5163 C	PUR	PUROMYCIN
REGION	6005	5782 C	tK	tK promoter
REGION	6500	6006 C	Ori P	
GENE	8552	6500 C	EBNA -1	
REGION	8553	8752	sv40	



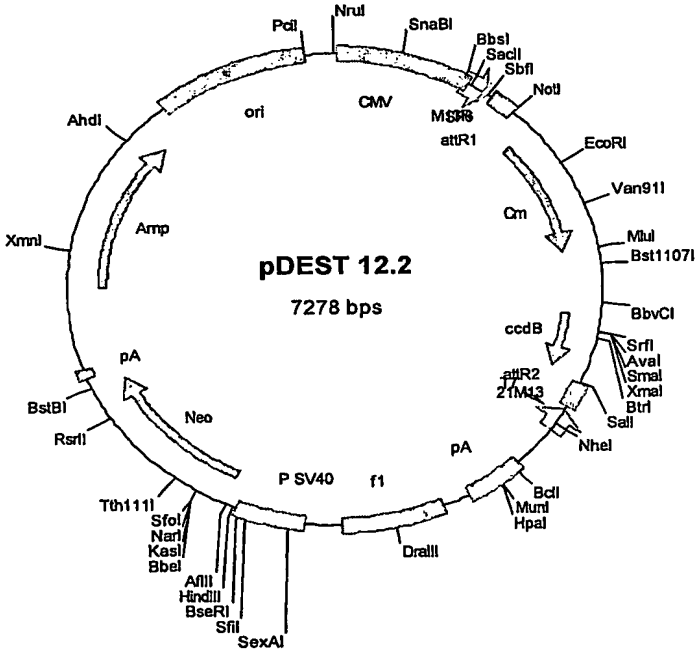
25

Figure 8:

Map of Expression vector pDEST12.2

Molecule: pDEST 12.2, 7278 bps DNA Circular
File Name: pDEST12-2.cm5
Description: Eukaryoric expression vector

Type	Start	End	Name	Description
REGION	15	608	CMV	CMV promoter
MARKER	648		M13R	M13R primer
REGION	687	706	SP6	SP6 promoter
REGION	730	854	at tR1	
GENE	963	1622	Cm	
GENE	1964	2269	ccdB	
REGION	2310	2434	attR2	
GENE	2484	2464 C	T7	T7 promoter
MARKER	2512	C	21M13	21M13 primer
REGION	2784	3050	p A	SV40 polyadenylation signal
REGION	3176	3631	f1	f1 intergenic region
REGION	3791	4099	P SV40	SV40 ori & early promoter
GENE	4158	4952	Neo	
REGION	5016	5064	pA	synthetic poly adenylation signal
GENE	5475	6335	Amp	
REGION	6480	7153	ori	pUC ori



25

Figure 9

Map of pDONR 221

Molecule: pDONR221, 4759 bps DNA Circular
File Name: pDONR221.cm5, dated 03 Jun 2003
Description:

Type	Start	End	Name	Description
REGION	295	268 C	rrnB T2	transcription termination sequence
REGION	470	427 C	rrnB T1	transcription termination sequence
REGION	536	553	21M13	M13 Forward primer
REGION	570	801	attP1	
GENE	1197	1502	ccdB	
GENE	1844	2503	Cm r	Chloramphenicol resistance gene
REGION	2751	2982	attP2	
REGION	3040	3023 C	M13 Rev	M13 Reverse primer
GENE	3153	3962	Kan r	
REGION	4083	4756	pUC ori	

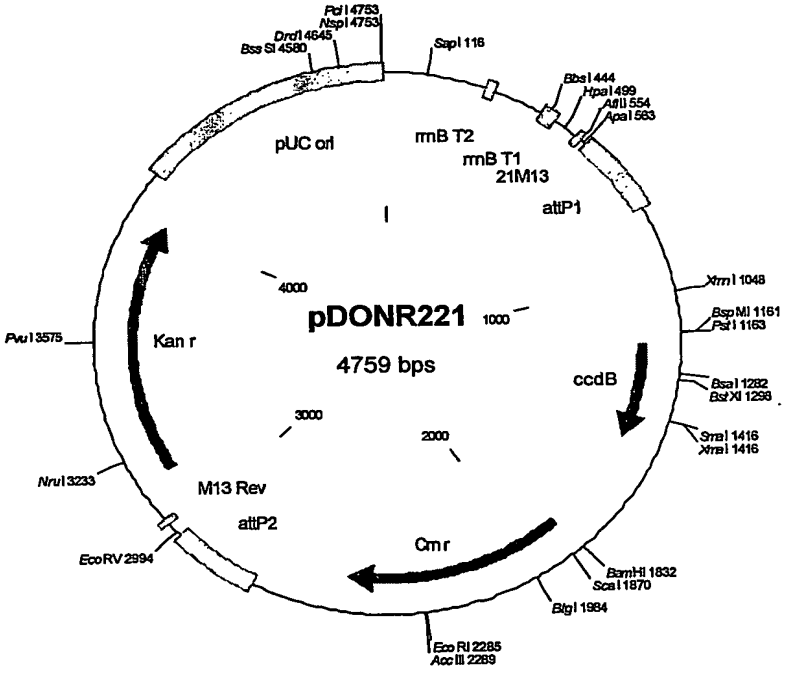
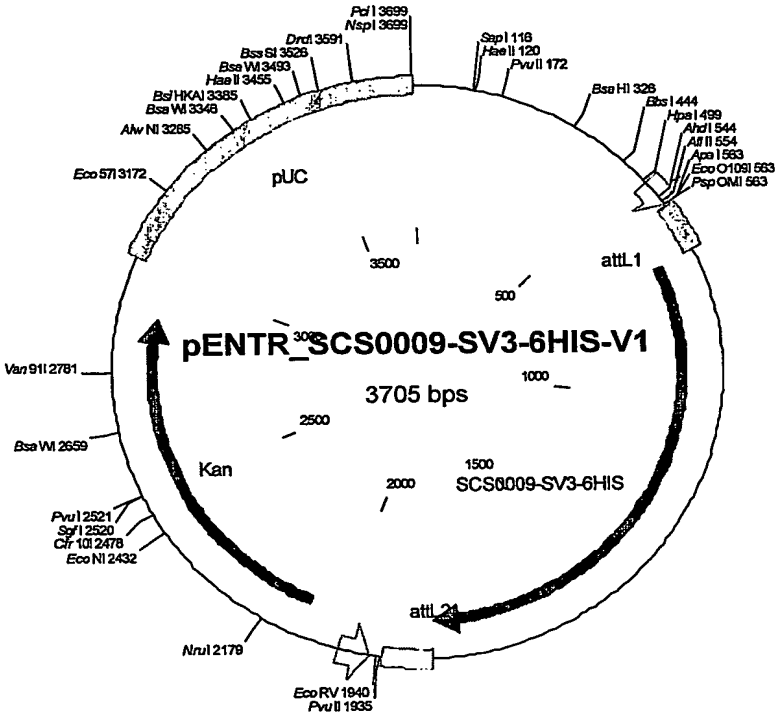


Figure 10

Map of pENTR-SCS0009SV3-6HIS

Molecule: pENTR_SCS0009 -SV3-6HIS-V1, 3705 bps DNA Circular
File Name: 14879[1].cm5
Description: Ligation of SCS0009 -SV3-6HIS into pdonr221

Type	Start	End	Name	Description
MARKER	536			21M13 Forward primer
REGION	568	670	attL1	
GENE	677	1825	SCS0009 -SV3-6HIS	
REGION	1829	1930	attL2	
MARKER	1986	C		M13 Reverse primer
GENE	2099	2908	Kan	Kanamycin resistance gene
REGION	3029	3702	pUC	pUC origin



15

Figure 11

Map of pEAK12d-SCS0009SV3-6HIS

Molecule: pEAK12d-SCS0009-SV3-6HIS-V1, 8098 bps DNA Circula r
File Name: 14885[1].cm5
Description: pEAK12 DES with two recombination sites attR1 and attR2 between which the cDNA is inserted

Type	Start	End	Name	Description
REGION	2	595	pmb -ori	
GENE	596	1519	Amp	
REGION	1690	2795	EF -1alpha	
REGION	2796	2845		MCS''
REGION	2855	2874	attB1	
GENE	2888	4039	SCS0009 -SV3-6HIS	
REGION	4044	4065	attB2	
REGION	4071	4071		'MCS
REGION	4072	4500	'A	poly A/splice
GENE	5119	4501 C	PUR	PUROMYCIN
REGION	5343	5120 C	tK	tK promoter
REGION	5838	5344 C	Ori P	
GENE	7890	5838 C	EBNA -1	
REGION	7891	8090	sv 40	

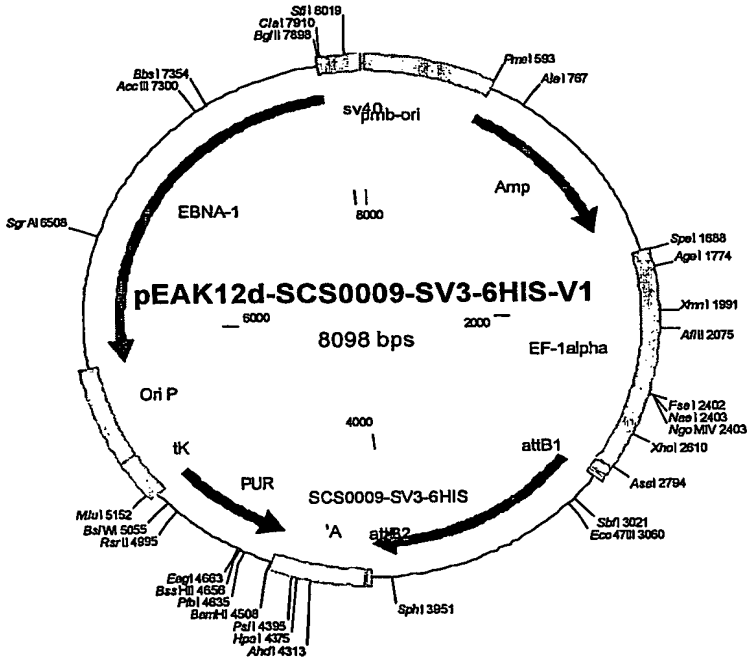


Figure 12

Map of pDEST12.2-SCS0009SV3-6HIS
Molecule: pDEST12.2_SCS0009-SV3-6HIS-V1, 6784 bps DNA Circular
File Name: 14889[1].cm5
Description: Ligation of SCS0009-SV3-6HIS-V1 into pDEST 12.2

Type	Start	End	Name	Description
REGION	15	608	CMV promoter	
REGION	648	665	M13R	Forward primer
REGION	687	706	SP6 promoter	
REGION	730	756	attB1	
GENE	763	1911	SCS0009-SV3-6HIS	
REGION	1940	1915 C	attB2	
GENE	1990	1970 C	T7	T7 promoter
REGION	2018	2001 C		21M13 reverse primer
REGION	2290	2556	pA	SV40 polyadenylation signal
REGION	2682	3137	f1	f1 intergenic region
REGION	3297	3605	P SV40	SV40 ori & early promoter
GENE	3664	4458	Neo	
REGION	4522	4570	pA	synthetic poly adenylation signal
GENE	4981	5841	Amp	
REGION	5986	6659	ori	pUC ori

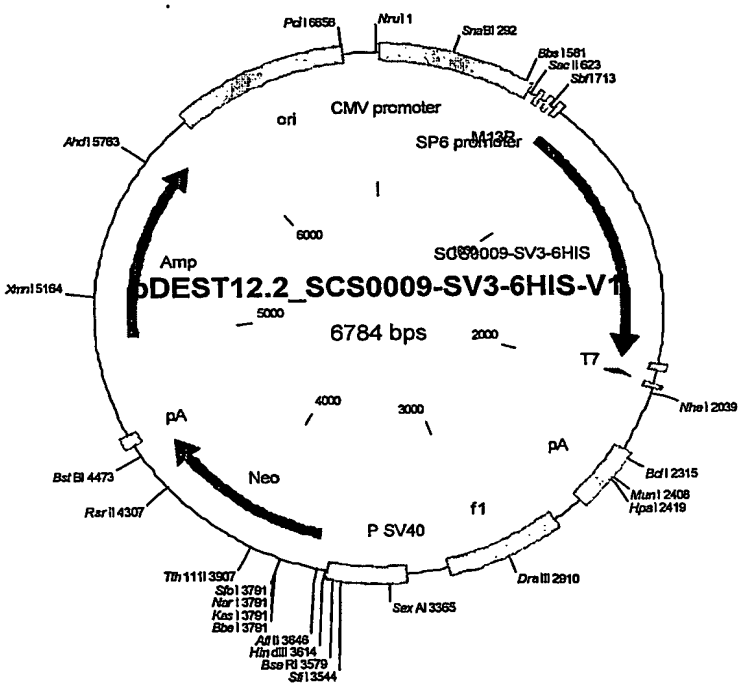


Figure 13

Nucleotide sequence of SCS0009 prediction with translation

```
   1  AGACGGCAAC GTGGACAGGA AGAAGCGGAG GGCGAGGAGG AGCAGAGGAG CACACAGATG
  61  AAGCAGGTGT CCACGCGTCC GGCCGTCCAT CCGTCCGTCC CTCCTGGGGC CGGCGCTGAC
 121  CATGCCCAGC GGCTGCCGCT GCCTGCATCT  CGTGTGCCTG TTGTGCATTC TGGGGGCTCC
        M  P  S   G  C  R   C  L  H   L  V  C  L   L  C  I   L  G  A

 181  CGGTCAGCCT GTCCGAGCCG ATGACTGCAG CTCCCACTGT GACCTGGCCC ACGGCTGCTG
      P  G  Q  P   V  R  A   D  D  C   S  S  H  C   D  L  A   H  G  C

 241  TGCACCTGAC GGCTCCTGCA GGTGTGACCC GGGCTGGGAG GGGCTGCACT GTGAGCGCTG
      C  A  P  D   G  S  C   R  C  D   P  G  W  E   G  L  H   C  E  R

 301  TGTGAGGATG CCTGGCTGCC AGCACGGTAC CTGCCACCAG CCATGGCAGT GCATCTGCCA
      C  V  R  M   P  G  C   Q  H  G   T  C  H  Q   P  W  Q   C  I  C

 361  CAGTGGCTGG GCAGGCAAGT TCTGTGACAA AGGCTTCCAT GGGCGTGACT GCGAGCGCAA
      H  S  G  W   A  G  K   F  C  D   K  G  F  H   G  R  D   C  E  R

 421  GGCTGGACCC TGTGAACAGG CAGGCTCCCC ATGCCGCAAT GGCGGGCAGT GCCAGGACGA
      K  A  G  P   C  E  Q   A  G  S   P  C  R  N   G  G  Q   C  Q  D

 481  CCAGGGCTTT GCTCTCAACT TCACGTGCCG CTGCTTGGTG GGCTTTGTGG GTGCCCGCTG
      D  Q  G  F   A  L  N   F  T  C   R  C  L  V   G  F  V   G  A  R

 541  TGAGGTAAAT GTGGATGACT GCCTGATGCG GCCTTGTGC T AACGGTGCCA CCTGCCTTGA
      C  E  V  N   V  D  D   C  L  M   R  P  C  A   N  G  A   T  C  L

 601  CGGCATAAAC CGCTTCTCCT GCCTCTGTCC TGAGGGCTTT GCTGGACGCT TCTGCACCAT
      D  G  I  N   R  F  S   C  L  C   P  E  G  F   A  G  R   F  C  T

 661  CAACCTGG AT GACTGTGCCA GCCGCCCATG CCAGAGAGGG GCCCGCTGTC GGGACCGTGT
      I  N  L  D   D  C  A   S  R  P   C  Q  R  G   A  R  C   R  D  R

 721  CCACGACTTC GACTGCCTCT GCCCCAGTGG CTATGGTGGC AAGACCTGTG AGCTTGTCTT
      V  H  D  F   D  C  L   C  P  S   G  Y  G  G    K  T  C   E  L  V

 781  ACCTGTCCCA GACCCCCCAA CCACAGTGGA CACCCCTCTA GGGCCCACCT CAGCTGTAGT
      L  P  V  P   D  P  P   T  T  V   D  T  P  L   G  P  T   S  A  V

 841  GGTACCTGCC ACGGGGCCAG CCCCCCACAG CGCAGGGGCT GGTCTGCTGC GGATCTCAGT
      V  V  P  A   T  G  P   A  P  H   S  A  G  A   G  L  L   R  I  S

 901  GAAGGAGGTG GTGCGGAGGC AAGAGGCTGG GCTAGGTGAG CCTAGCTTGG TGGCCCTGGT
      V  K  E  V   V  R  R   Q  E  A   G  L  G  E   P  S  L   V  A  L

 961  GGTGTTTGGG GCCCTCACTG CTGCCCTGGT TCTGGCTACT GTGTTGC TGA CCCTGAGGGC
      V  V  F  G   A  L  T   A  A  L   V  L  A  T   V  L  L   T  L  R

1021  CTGGCGCCGG GGTGTCTGCC CCCCTGGACC CTGTTGCTAC CCTGCCCCAC ACTATGCTCC
      A  W  R  R   G  V  C   P  P  G   P  C  C  Y   P  A  P   H  Y  A

1081  AGCGTGCCAG GACCAG GAGT GTCAGGTTAG CATGCTGCCA GCAGGGCTCC CCCTGCCACG
      P  A  C  Q   D  Q  E   C  Q  V   S  M  L  P   A  G  L   P  L  P

1141  TGACTTGCCC CCTGAGCCTG GAAAGACCAC AGCACTGTGA TGGAGGTGGG GGCTTTCTGG
      R  D  L  P   P  E  P   G  K  T   T  A  L

1201  CCCCCT TCCT CACCTCTTCC ACCCCTCAGA CTGGAGTGGT CCGTTCTCAC CACCCTTCAG
1261  CTTGGGTACA CACACAGAGG AGACCTCAGC CTCACACCAG AAATATTATT TTTTTAATAC
```

```
1321  ACAGAATGTA AGATGGAATT TTATCAAATA AAACTATGAA AATGCAAGTG GGCTCCTATG
1381  CCAGAAAAAC CCACCTGGCG TTCCAGATGC AAGAGGGCCA  GAGCAGAGGC CTGGTTCTGG
1441  GGAAGCCTCA GGATGCTGCC CACCAAGGAG TGATTTCCAA AGAGTAATCC AGGGTGCCCT
1501  TTTCCCTTCT GGGGAAGTGT GGAGAGGTAG AGCCCCAGAG GAGAATGTAA ACAAGCAGCC
1561  AGCACCTCTG TATAGGCCCG GCCTGGATCA GAGAGAGGGG AGAACTCTGC AGGGTGTGGG
1621  ATTGGGC TCA GGGACCTCCG AGTGAGGCAG GGACTCCCTG CTG
```

Figure 14

Nucleotide sequence with translation of cDNA insert in image clone 3349698 (SCS0009 -SV4)

```
  1  ATGCCCAGCG GCTGCCGCTG CCTGCATCTC GTGTGCCTGT TGTGCATTCT GGGGGCTCCC
       M  P  S   G  C  R   C  L  H   L   V  C  L   L  C  I   L  G  A  P

 61  GGTCAGCCTG TCCGAGCCGA TGACTGCAGC TCCCACTGTG ACCTGGCCCA CGGCTGCTGT
       G  Q  P   V  R  A   D  D  C   S   S  H  C   D  L  A   H  G  C  C

121  GCACCTGACG GCTCCTGCAG GTGTGACCCG GGCTGGGAGG GGCTGCACTG TGAGCGCTGT
       A  P  D   G  S  C   R  C  D   P   G  W  E   G  L  H   C  E  R  C

181  GTGAGGATGC CTGGCTGCCA GCACGGTACC TGCCACCAGC CATGGCAGTG CATCTGCCAC
       V  R  M   P  G  C   Q  H  G   T   C  H  Q   P  W  Q   C  I  C  H

241  AGTGGCTGGG CAGGCAAGTT CTGTGACAAA G ATGAACATA TCTGTACCAC GCAGTCCCCC
       S  G  W   A  G  K   F  C  D   K   D  E  H   I  C  T   T  Q  S  P

301  TGCCAGAATG GAGGCCAGTG CATGTATGAC GGGGGCGGTG AGTACCATTG TGTGTGCTTA
       C  Q  N   G  G  Q   C  M  Y   D   G  G  G   E  Y  H   C  V  C  L

361  CCAGGCTTCC ATGGGCGTGA CTGCGAGCGC AAGGCTGGAC CC
       P  G  F   H  G  R   D  C  E   R   K  A  G   P
```

Figure 15

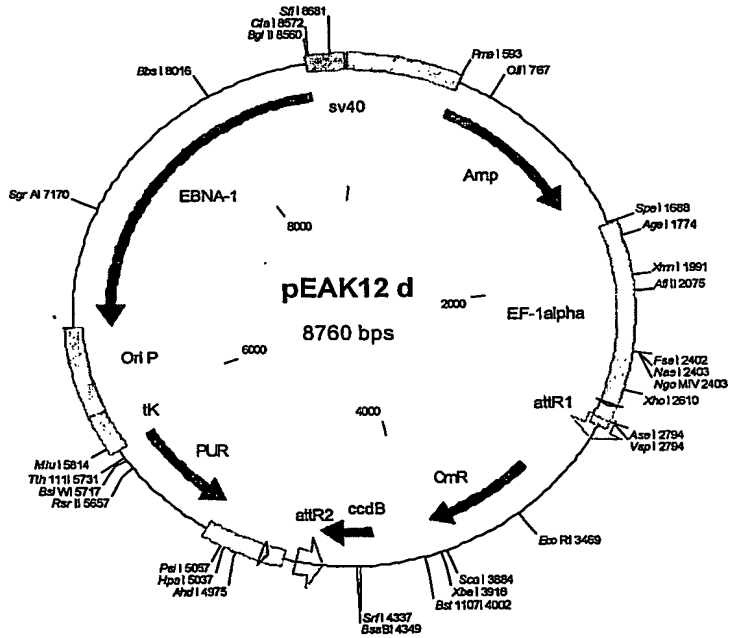
Alignment of predicted amino acid sequence of SCS0009 with SCS0009-SV4

```
SCS0009      MPSGCRCLHLVCLLCILGAPGQPVRADDCSSHCDLAHGCCAPDGSCRCDP
SCS0009SV4   MPSGCRCLHLVCLLCILGAPGQPVRADDCSSHCDLAHGCCAPDGSCRCDP

SCS0009      GWEGLHCERCVRMPGCQHGTCHQPWQCICHSGWAGKFCDK ----------
SCS0009SV4   GWEGLHCERCVRMPGCQHGTCHQPWQCICHSGWAGKFCDKDEHICTTQSP

SCS0009      ---------------------GFHGRDCERKAGPCEQAGSPCRNGGQCQD
SCS0009SV4   CQNGGQCMYDGGGEYHCVCLPGFHGRDCERKAGP ----------------

SCS0009      DQGFALNFTCRCLVGFVGARCEVNVDDCLMRPCANGATCLDGINRFSCLC
SCS0009SV4   --------------------------------------------------

SCS0009      PEGFAGRFCTINLDDCASRPCQRGARCRDRVHDFDCLCPSGYGGKTCELV
SCS0009SV4   --------------------------------------------------

SCS0009      LPVPDPPTTVDTPLGPTSAVVVPATGPAPHSAGAGLLRISVKEVVRRQEA
SCS0009SV4   --------------------------------------------------

SCS0009      GLGEPSLVALVVFGALTAALVLATVLLTLRAWRRGVCPPGPCCYPAPHYA
SCS0009SV4   --------------------------------------------------

SCS0009      PACQDQECQVSMLPAGLPLPRDLPPEPGKTTAL
SCS0009SV4   ---------------------------------
```

Figure 16

Map of expression vector pEAK12d
Molecule: pEAK12 d, 8760 bps DNA Circular
File Name: pEAK12DEST.cm5
Description: Mammalian cell expression vector (plasmid ID 11345)
Molecule Features:

Type	Start	End	Name	Description
REGION	2	595		pmb -ori
GENE	596	1519	Amp	
REGION	1690	2795	EF-1alpha	
REGION	2703	2722		position of pEAK12F primer
REGION	2796	2845		MCS
MARKER	2855		attR1	
GENE	3256	3915	CmR	
GENE	4257	4562	ccdB	
MARKER	4603	C	attR2	
REGION	4733	4733		MCS
REGION	4734	5162		poly A/splice
REGION	4819	4848 C		position of pEAK12R primer
GENE	5781	5163 C	PUR	PUROMYCIN
REGION	6005	5782 C	tK	tK promoter
REGION	6500	6006 C	Ori P	
GENE	8552	6500 C	EBNA -1	
REGION	8553	8752	sv40	



25

Figure 17

Map of Expression vector pDEST12.2

Molecule: pDEST 12.2, 7278 bps DNA Circular
File Name: pDEST12-2.cm5
Description: Eukaryoric expression vector

Type	Start	End	Name	Description
REGION	15	608	CMV	CMV promoter
MARKER	648		M13R	M13R primer
REGION	687	706	SP6	SP6 promoter
REGION	730	854	attR1	
GENE	963	1622	Cm	
GENE	1964	2269	ccdB	
REGION	2310	2434	attR2	
GENE	2484	2464 C	T7	T7 promoter
MARKER	2512	C	21M13	21M13 primer
REGION	2784	3050	pA	SV40 polyadenylation signal
REGION	3176	3631	f1	f1 intergenic region
REGION	3791	4099	P SV40	SV40 ori & early promoter
GENE	4158	4952	Neo	
REGION	5016	5064	pA	synthetic poly adenylation signal
GENE	5475	6335	Amp	
REGION	6480	7153	ori	pUC ori

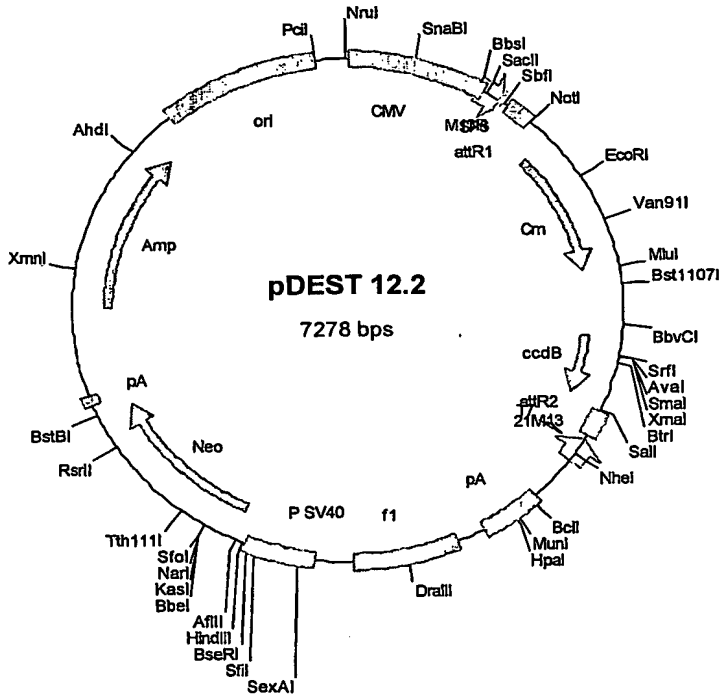


Figure 18

Map of pDONR 221

Molecule: pDONR221, 4759 bps DNA Circular
File Name: pDONR221.cm5 , dated 03 Jun 2003
Description:

Type	Start	End	Name	Description
REGION	295	268 C	rrnB T2	transcription termination sequence
REGION	470	427 C	rrnB T1	transcription termination sequence
REGION	536	553	21M13	M13 Forward primer
REGION	570	801	attP1	
GENE	1197	1502	ccdB	
GENE	1844	2503	Cm r	Chloramphenicol resistance gene
REGION	2751	2982	attP2	
REGION	3040	3023 C	M13 Rev	M13 Reverse primer
GENE	3153	3962	Kan r	
REGION	4083	4756	pUC ori	

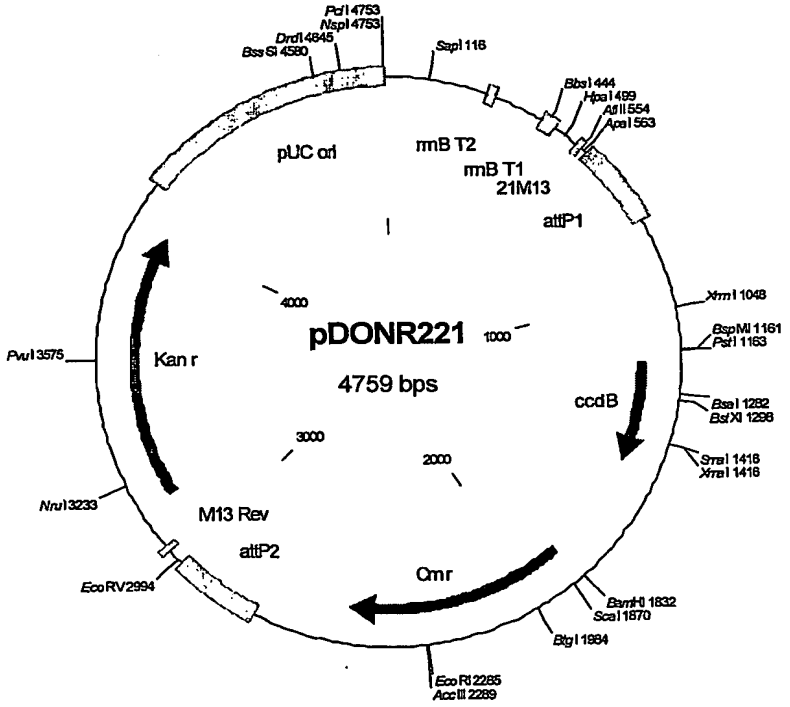
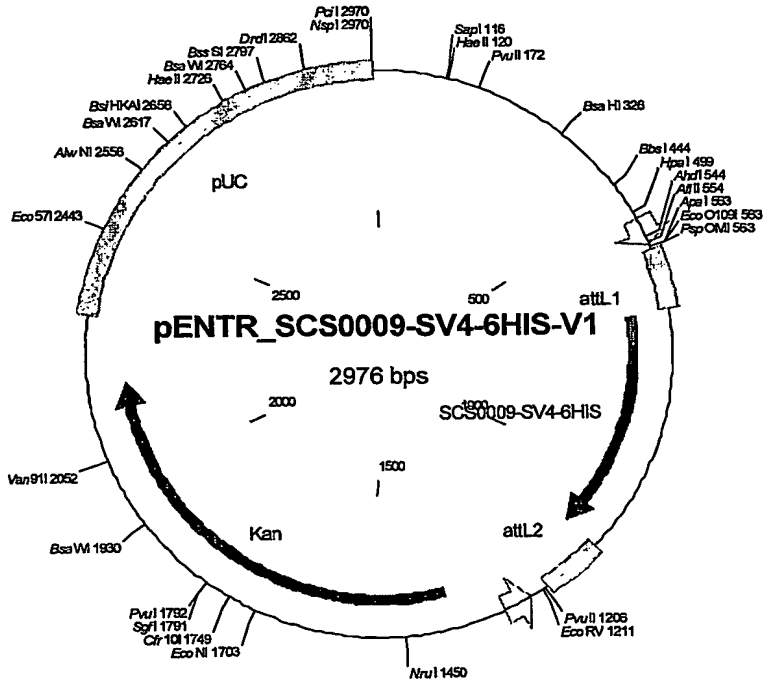


Figure 19

Map of pENTR-SCS0009SV4-6HIS

Molecule: pENTR_SCS0009-SV4-6HIS-V1, 2976 bps DNA Circular
File Name: 15055[1].cm5
Description: Ligation of SCS0009-SV4-6HIS into pdonr221

Type	Start	End	Name	Description
MARKER	536			21M13 Forward primer
REGION	568	670	attL1	
GENE	677	1096	SCS0009-SV4-6HIS	
REGION	1100	1201	attL2	
MARKER	1257	C		M13 Reverse primer
GENE	1370	2179	Kan	Kanamycin resistance gene
REGION	2300	2973	pUC	pUC origin



15

Figure 20

Map of pEAK12d-SCS0009SV4-6HIS

Molecule: pEAK12d-SCS0009-SV4-6HIS-V1, 7369 bps DNA Circular
File Name: 15061[1].cm5
Description: pEAK12 DES with two recombination sites attR1 and attR2 between which the cDNA is inserted

Type	Start	End	Name	Description
REGION	2	595	pmb-ori	
GENE	596	1519	Amp	
REGION	1690	2795	EF-1alpha	
REGION	2796	2845		MCS''
REGION	2855	2874		attB1
GENE	2888	3310	SCS0009-SV4-6HIS	
REGION	3315	3336		attB2
REGION	3342	3342		'MCS
REGION	3343	3771	'A	poly A/splice
GENE	4390	3772 C	PUR	PUROMYCIN
REGION	4614	4391 C	tK	tK promoter
REGION	5109	4615 C	Ori P	
GENE	7161	5109 C	EBNA-1	
REGION	7162	7361	sv40	

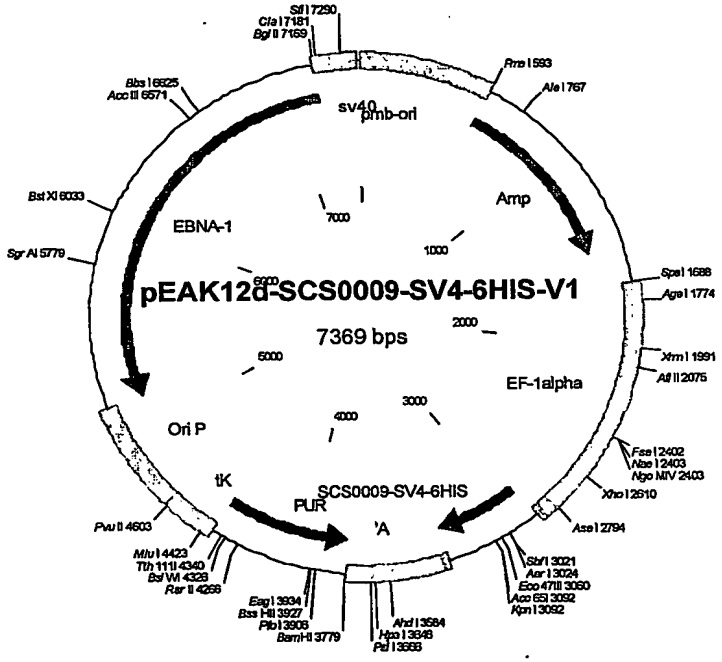


Figure 21
Map of pDEST12.2-SCS0009SV4-6HIS

Molecule: pDEST12.2_SCS0009 -SV4-6HIS-V1, 6055 bps DNA Circular
File Name: 15063[1].cm5
Description: Ligation of SCS0009 -SV4-6HIS-V1 into pDEST 12.2

Type	Start	End	Name	Description
REGION	15	608	CMV promoter	
REGION	648	665	M13R	Forward primer
REGION	687	706	SP6 promoter	
REGION	730	756		attB1
GENE	763	1182	SCS0009 -SV4-6HIS	
REGION	1211	1186 C		attB2
GENE	1261	1241 C	T7	T7 promoter
REGION	1289	1272 C		21M13 reverse primer
REGION	1561	1827	pA	SV40 polyadenylation signal
REGION	1953	2408	f1	f1 intergenic region
REGION	2568	2876	P SV40	SV40 ori & early promoter
GENE	2935	3729	Neo	
REGION	3793	3841	pA	synthetic poly adenylation signal
GENE	4252	5112	Amp	
REGION	5257	5930	ori	pUC ori

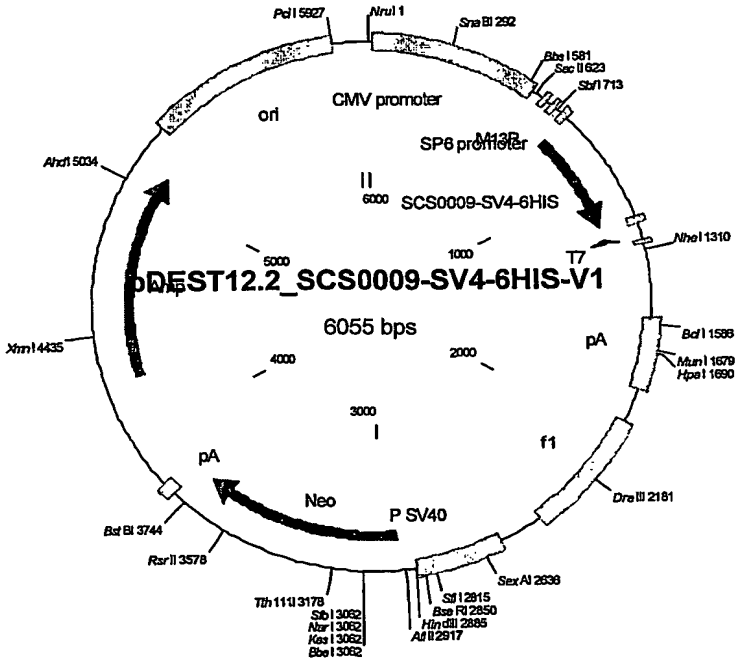


Figure 22

Nucleotide sequence of SCS0009 prediction with translation

```
   1  agacggcaac gtggacagga agaagcggag ggcgaggagg agcagaggag cacacagatg
  61  aaqcaggtgt ccacgcgtcc ggccgtccat ccgtccgtcc ct cctggggc cggcgctgac
 121  catgcccagc ggctgccgct gcctgcatct cgtgtgcctg ttgtgcattc tgggggctcc
         m  p  s   g  c  r   c  l  h   l  v  c  l   l  c  i   l  g  a

 181  cggtcagcct gtccgagccg atgactgcag ctcccactgt gacctggccc acggctgctg
      p  g  q  p   v  r  a   d  d  c   s  s  h  c   d  l  a   h  g  c

 241  tgcacctgac ggctcctgca ggtgtgaccc gggctgggag gggctgcact gtgagcgctg
      c  a  p  d   g  s  c   r  c  d   p  g  w  e   g  l  h   c  e  r

 301  tgtgaggatg cctggctgcc agcacggtac ctgccaccag ccatgg cagt gcatctgcca
      c  v  r  m   p  g  c   q  h  g   t  c  h  q   p  w  q   c  i  c

 361  cagtggctgg gcaggcaagt tctgtgacaa aggcttccat gggcgtgact gcgagcgcaa
      h  s  g  w   a  g  k   f  c  d   k  g  f  h   g  r  d   c  e  r

 421  ggctggaccc tgtgaacagg caggctcccc atgccgcaat ggcgggcagt gccaggacga
      k  a  g  p   c  e  q   a  g  s   p  c  r  n   g  g  q   c  q  d

 481  ccagggcttt gctctcaact tcacgtgccg ctgcttggtg ggctttgtgg gtgcccgctg
      d  q  g  f   a  l  n   f  t  c   r  c  l  v   g  f  v   g  a  r

 541  tgaggtaaat gtggatgact gcctgatgcg gccttgtgct aacggtgcca cctgccttga
      c  e  v  n   v  d  d   c  l  m   r  p  c  a   n  g  a   t  c  l

 601  cggcataaac cgcttctcct gcctctgtcc tgagggcttt gctggacgct tctgcaccat
      d  g  i  n   r  f  s   c  l  c   p  e  g  f   a  g  r   f  c  t

 661  caacctggat gactgtgcca gccgcccatg ccagagaggg gcccgctgtc gggaccgtgt
      i  n  l  d   d  c  a   s  r  p   c  q  r  g   a  r  c   r  d  r

 721  ccacgacttc gactgcctct gccccagtgg ctatggtggc aagacctgtg agct tgtctt
      v  h  d  f   d  c  l   c  p  s   g  y  g  g   k  t  c   e  l  v

 781  acctgtccca gacccccaa ccacagtgga cacccctcta gggcccacct cagctgtagt
      l  p  v  p   d  p  p   t  t  v   d  t  p  l   g  p  t   s  a  v

 841  ggtacctgcc acggggccag ccc cccacag cgcaggggct ggtctgctgc ggatctcagt
      v  v  p  a   t  g  p   a  p  h   s  a  g  a   g  l  l   r  i  s

 901  gaaggaggtg gtgcggaggc aagaggctgg gctaggtgag cctagcttgg tggccctggt
      v  k  e  v   v  r  r   q  e  a   g  l  g  e   p  s  l   v  a  l

 961  ggtgtttggg gccctcactg ctgccctggt tctggctact gtgttgctga ccctgagggc
      v  v  f  g   a  l  t   a  a  l   v  l  a  t   v  l  l   t  l  r

1021  ctggcgccgg ggtgtctgcc cccctggacc ctgttgctac cctgccccac actatgctcc
      a  w  r  r   g  v  c   p  p  g   p  c  c  y   p  a  p   h  y  a

1081  agcgtgccag gaccaggagt gtcaggttag catgctgcca gcagggctcc ccctgccacg
      p  a  c  q   d  q  e   c  q  v   s  m  l  p   a  g  l   p  l  p
```

```
1141  tgacttgccc cctgagcctg gaaagaccac agcactgtga tggaggtggg ggctttctgg
      r  d  l  p    p  e  p    g  k  t    t  a  l

1201  ccccttcct cacctcttcc acccctcaga ctggagtggt ccgttctcac cacccttcag
1261  cttgggtaca cacacagagg agacctcagc ctcacaccag aaatattatt tttttaatac
1321  acagaatgta agatggaatt ttatcaaata aaactatgaa aatgcaagtg g gctcctatg
1381  ccagaaaaac ccacctggcg ttccagatgc aagagggcca gagcagaggc ctggttctgg
1441  ggaagcctca ggatgctgcc caccaaggag tgatttccaa agagtaatcc agggtgccct
1501  tttcccttct ggggaagtgt ggagaggtag agccccagag gagaatgtaa acaagcagcc
1561  agcacctctg tataggccc g gcctggatca gagagagggg agaactctgc agggtgtggg
1621  attgggctca gggacctccg agtgaggcag ggactccctg ctg
```

Figure 23

Nucleotide sequence with translation of SCS0009-SV5 PCR product indicating the positions of the SCS0009-AP1, -AP2, -AP3 and –AP4 primers used to generate the SCS0009 sequence.

```
                                                          SCS0009-AP1
  1  tccatccgtc cgtccctcct ggggccggcg ctgaccatgc ccagcggctg ccgctgcctg
                                                m  p  s  g   c  r  c  l

 61  catctcgtgt gcctgttgtg cattctgggg gctcccggtc agcctgtccg agc cgatgac
      h  l  v   c  l  l   c  i  l  g   a  p  g   q  p  v   r  a  d  d

121  tgcagctccc actgtgacct ggcccacggc tgctgtgcac ctgacggctc ctgcaggtgt
      c  s  s   h  c  d   l  a  h  g   c  c  a   p  d  g   s  c  r  c

181  gacccgggct gggaggggct gcactgtgag cgctgtgtga ggatgcctgg ctgccagcac
      d  p  g   w  e  g   l  h  c  e   r  c  v   r  m  p   g  c  q  h

241  ggtacctgcc accagccatg gcagtgcatc tgccacagtg gctgggcagg caagttctgt
      g  t  c   h  q  p   w  q  c  i   c  h  s   g  w  a   g  k  f  c
                                                           SCS0009-AP2
301  gacaaagatg aacatatctg taccacgcag tccccctgcc agaatggagg ccagtgcatg
      d  k  d   e  h  i   c  t  t  q   s  p  c   q  n  g   g  q  c  m
                                                      SCS0009-AP3
361  tatgacgggg gcggtgagta ccattgtgtg tgcttaccag gcttccatgg gcgtgactgc
      y  d  g   g  g  e   y  h  c  v   c  l  p   g  f  h   g  r  d  c

421  gagcgcaagg ctggaccctg tgaacaggca ggctccccat gccgcaatgg cgggcagtgc
      e  r  k   a  g  p   c  e  q  a   g  s  p   c  r  n   g  g  q  c

481  caggacgacc agggctttgc tctcaacttc acgtgccgct gcttggtggg ctttgtgggt
      q  d  d   q  g  f   a  l  n  f   t  c  r   c  l  v   g  f  v  g

541  gcccgctgtg aggtaaatgt ggatgactgc ctgatgcggc cttgtgctaa cggtgccacc
      a  r  c   e  v  n   v  d  d  c   l  m  r   p  c  a   n  g  a  t

601  tgccttgacg gcataaaccg cttctcctgc ctctgtcctg agggctttgc tggacgcttc
      c  l  d   g  i  n   r  f  s  c   l  c  p   e  g  f   a  g  r  f

661  tgcaccatca acctggatga ctgtgccagc cgcccatgcc agagaggggc ccgctgtcgg
      c  t  i   n  l  d   d  c  a  s   r  p  c   q  r  g   a  r  c  r

721  gaccgtgtcc acgacttcga ctgcctctgc cccagtggct atggtggcaa gacctgtgag
      d  r  v   h  d  f   d  c  l  c   p  s  g   y  g  g   k  t  c  e

781  cttgtcttac ctgtcccaga ccccccaacc acagtggaca cccctctagg gcccacctca
      l  v  l   p  v  p   d  p  p  t   t  v  d   t  p  l   g  p  t  s
```

```
841  gctgtagtgg tacctgccac ggggccagcc ccccacagcg caggggctgg tctgctgcgg
       a  v  v  v  p  a  t  g  p  a  p  h  s  a  g  a  g  l  l  r

901  atctcagtga aggaggtggt gcggaggcaa gagg ctgggc taggtgagcc tagcttggtg
       i  s  v  k  e  v  v  r  r  q  e  a  g  l  g  e  p  s  l  v

961  gccctggtgg tgtttggggc cctcactgct gccctggttc tggctactgt gttgctgacc
       a  l  v  v  f  g  a  l  t  a  a  l  v  l  a  t  v  l  l  t

1021 ctgagggcct ggcgccgggg tgtctgcccc cctggaccct gttgctaccc tgccccacac
       l  r  a  w  r  r  g  v  c  p  p  g  p  c  c  y  p  a  p  h

1081 tatgctccag cgtgccagga ccaggagtgt caggttagca tgctgccagc agggctcccc
       y  a  p  a  c  q  d  q  e  c  q  v  s  m  l  p  a  g  l  p

1141 ctgccacgtg acttgccccc tgagcctgga aagaccacag cactgtgatg gaggtgggg
       l  p  r  d  l  p  p  e  p  g  k  t  t  a  l
                <------------------------------------
                          SCS0009-AP4
```

Sequence in grey = bases not present in SCS0009 prediction

Figure 24

Nucleotide sequence and translation of cloned SCS0009 ORF

```
   1  accatgccca gcggctgccg ctgcctgcat ctcgtgtgcc tgttgtgcat tctgggggct
         m  p  s  g  c  r  c  l  h  l  v  c  l  l  c  i  l  g  a

  61  cccggtcagc ctgtccgagc cgatgactgc agctcccact gtgacctggc ccacggct gc
       p  g  q  p  v  r  a  d  d  c  s  s  h  c  d  l  a  h  g  c

 121  tgtgcacctg acggctcctg caggtgtgac ccgggctggg aggggctgca ctgtgagcgc
       c  a  p  d  g  s  c  r  c  d  p  g  w  e  g  l  h  c  e  r

 181  tgtgtgagga tgcctggctg cca gcacggt acctgccacc agccatggca gtgcatctgc
       c  v  r  m  p  g  c  q  h  g  t  c  h  q  p  w  q  c  i  c

 241  cacagtggct gggcaggcaa gttctgtgac aaaggcttcc atgggcgtga ctgcgagcgc
       h  s  g  w  a  g  k  f  c  d  k  g  f  h  g  r  d  c  e  r

 301  aaggctggac cctgtgaaca ggcaggctcc ccatgccgca atggcgggca gtgccaggac
       k  a  g  p  c  e  q  a  g  s  p  c  r  n  g  g  q  c  q  d

 361  gaccagggct ttgctctcaa cttcacgtgc cgctgcttgg tgggctttgt gggtgcccgc
       d  q  g  f  a  l  n  f  t  c  r  c  l  v  g  f  v  g  a  r

 421  tgtgaggtaa atgtggatga ctgcctgatg cggccttgtg ctaacggtgc cacctgcctt
       c  e  v  n  v  d  d  c  l  m  r  p  c  a  n  g  a  t  c  l

 481  gacggcataa accgcttctc ctgcctctgt cctgagggct ttgctggacg ctt ctgcacc
       d  g  i  n  r  f  s  c  l  c  p  e  g  f  a  g  r  f  c  t

 541  atcaacctgg atgactgtgc cagccgccca tgccagagag gggcccgctg tcgggaccgt
       i  n  l  d  d  c  a  s  r  p  c  q  r  g  a  r  c  r  d  r

 601  gtccacgact tcgactgcct ctgccccagt ggctatggtg gcaagacctg tgagcttgtc
       v  h  d  f  d  c  l  c  p  s  g  y  g  g  k  t  c  e  l  v

 661  ttacctgtcc cagacccccc aaccacagtg gacacccctc tagggcccac ctcagctgta
       l  p  v  p  d  p  p  t  t  v  d  t  p  l  g  p  t  s  a  v

 721  gtggtacctg ccacggggcc agcccccccac agcgcagggg ctggtctgct gcggatctca
       v  v  p  a  t  g  p  a  p  h  s  a  g  a  g  l  l  r  i  s

 781  gtgaaggagg tggtgcggag gcaagaggct gggctaggtg agcctagctt ggtggccctg
       v  k  e  v  v  r  r  q  e  a  g  l  g  e  p  s  l  v  a  l

 841  gtggtgtttg gggccctcac tgctgccctg gttctggcta ctgtgttgct gaccctgagg
       v  v  f  g  a  l  t  a  a  l  v  l  a  t  v  l  l  t  l  r

 901  gcctggcgcc ggggtgtctg ccccccctgga ccctgttgct accctgccc c acactatgct
       a  w  r  r  g  v  c  p  p  g  p  c  c  y  p  a  p  h  y  a

 961  ccagcgtgcc aggaccagga gtgtcaggtt agcatgctgc cagcagggct ccccctgcca
       p  a  c  q  d  q  e  c  q  v  s  m  l  p  a  g  l  p  l  p

1021  cgtgacttgc ccc ctgagcc tggaaagacc acagcactgt ga
       r  d  l  p  p  e  p  g  k  t  t  a  l
```

Figure 25

Map of pCR-BluntII-TOPO-SCS0009

Molecule: pCR-BluntII-TOPO-SCS0009, 5012 bps DNA Circular

Type	Start	End		Name	Description
GENE	1	336		LacZa'	LacZa gene
MARKER	239			SP6	SP6 promoter
MARKER	337			SCS0009 -AP4	SCS0009-AP4 primer site
MARKER	1107			SCS0009 -AP2	SCS0009-AP2 primer site
MARKER	1142		C	SCS0009 -AP3	SCS0009-AP3 primer site
GENE	1395	340	C	cds	SCS0009 cds
REGION	1398	337	C	Insert	SCS0009 assembly PCR product
MARKER	1398		C	SCS0009 -AP1	SCS0009-AP1 primer site
GENE	1399	1650		'LacZa	LacZa gene
MARKER	1487		C	T7	T7 promoter
REGION	1652	2066		f1 ori	
GENE	2400	3194		Kan	Kanamycin resistance ORF
GENE	3212	4072		Amp	Ampicillin resistance ORF
REGION	4217	4890		pUC ori	pUC origin

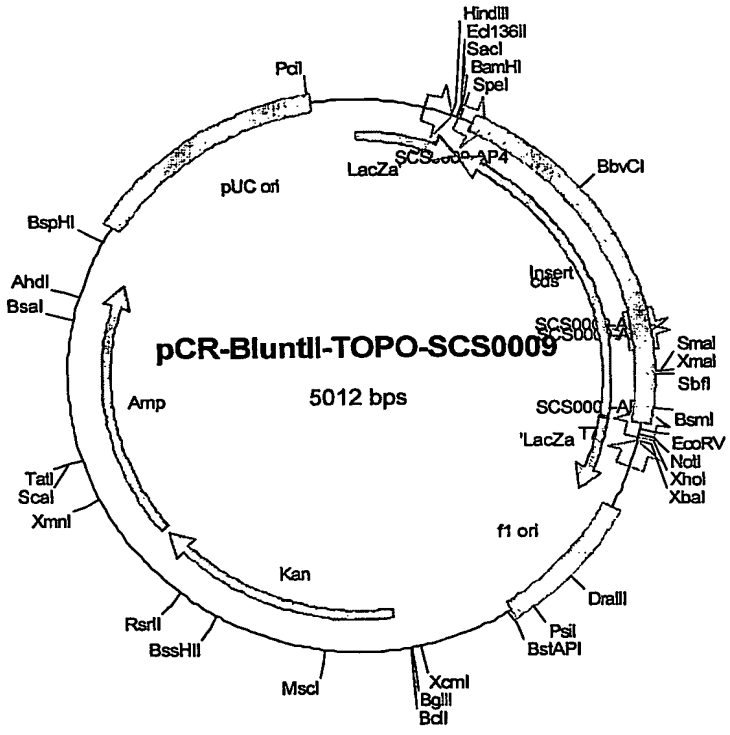


Figure 26

Map of pDONR 221

Molecule: pDONR221, 4759 bps DNA Circular
File Name: pDONR221.cm5
Description:

Type	Start	End	Name	Description
REGION	295	268 C	rrnB T2	transcription termination sequence
REGION	470	427 C	rrnB T1	transcription termination sequence
REGION	536	553	21M13	M13 Forward primer
REGION	570	801	attP1	
GENE	1197	1502	ccdB	
GENE	1844	2503	Cm r	Chloramphenicol resista nce gene
REGION	2751	2982	attP2	
REGION	3040	3023 C	M13 Rev	M13 Reverse primer
GENE	3153	3962	Kan r	
REGION	4083	4756	pUC ori	

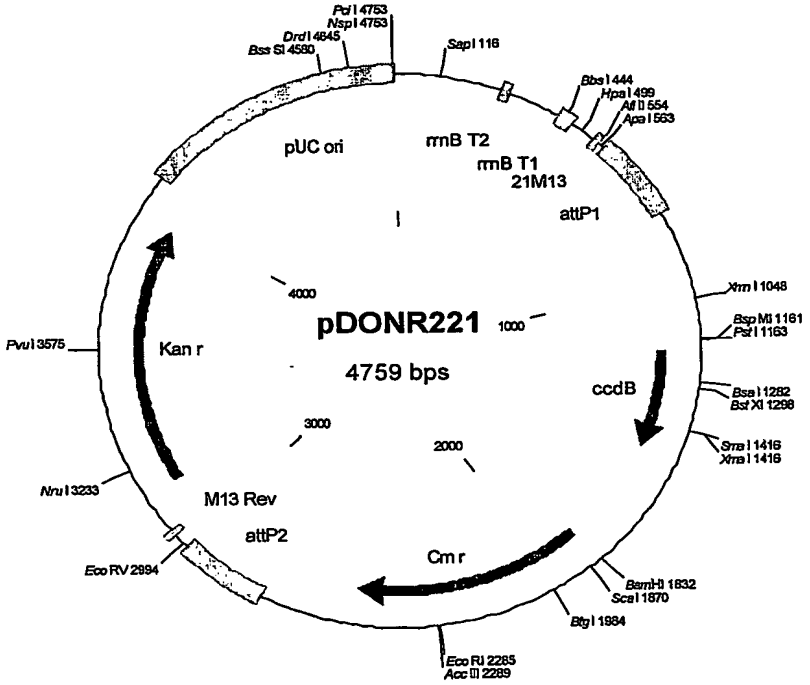


Figure 27

Map of expression vector pEAK12d

Molecule: pEAK12 d, 8760 bps DNA Circular
File Name: pEAK12DEST.cm5
Description: Mammalian cell expression vector (plasmid ID 11345)
Molecule Features:

Type	Start	End	Name	Description
REGION	2	595		pmb -ori
GENE	596	15 19	Amp	
REGION	1690	2795	EF -1alpha	
REGION	2703	2722		position of pEAK12F primer
REGION	2796	2845		MCS
MARKER	2855		attR1	
GENE	3256	3915	CmR	
GENE	4257	4562	ccdB	
MARKER	4603	C	attR2	
REGION	4733	4733		MCS
REGION	4734	5162		poly A/splice
REGION	4819	4848 C		position of pEAK12R primer
GENE	5781	5163 C	PUR	PUROMYCIN
REGION	6005	5782 C	tK	tK promoter
REGION	6500	6006 C	Ori P	
GENE	8552	6500 C	EBNA -1	
REGION	8553	8752	sv40	

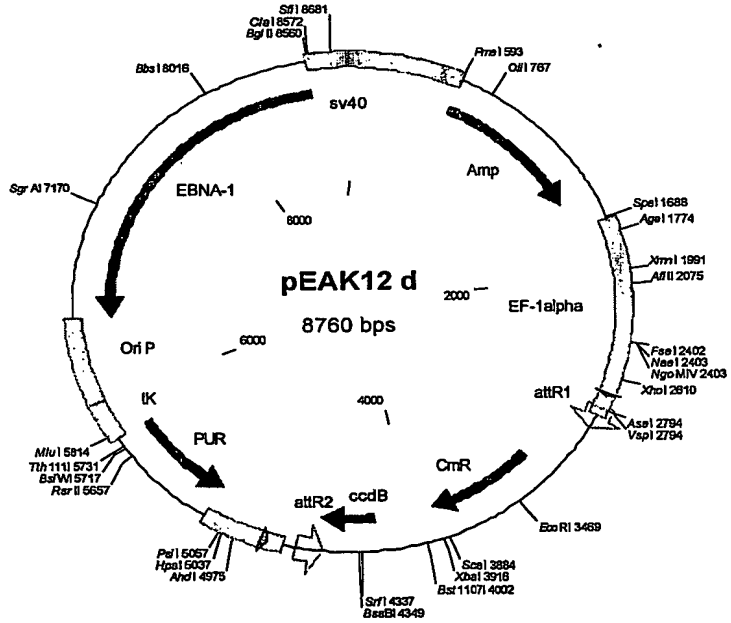


Figure 28

Map of Expression vector pDEST12.2

Molecule: pDEST 12 .2, 7278 bps DNA Circular
File Name: pDEST12 -2.cm5
Description: Eukaryoric expression vector

Type	Start	End	Name	Description
REGION	15	608	CMV	CMV promoter
MARKER	648		M13R	M13R primer
REGION	687	706	SP6	SP6 promoter
REGION	730	854	attR1	
GENE	963	1622	Cm	
GENE	1964	2269	ccdB	
REGION	2310	2434	attR2	
GENE	2484	2464 C	T7	T7 promoter
MARKER	2512	C	21M13	21M13 primer
REGION	2784	3050	pA	SV40 polyadenylation signal
REGION	3176	3631	f1	f1 intergenic region
REGION	3791	4099	P SV40	SV40 ori & early promoter
GENE	4158	4952	Neo	
REGION	5016	5064	pA	synthetic poly adenylation signal
GENE	5475	6335	Amp	
REGION	6480	7153	ori	pUC ori

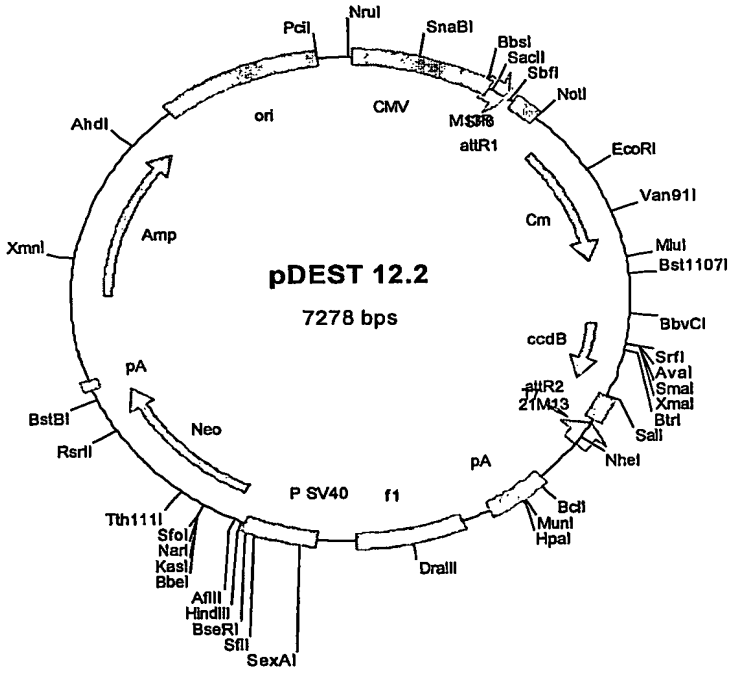
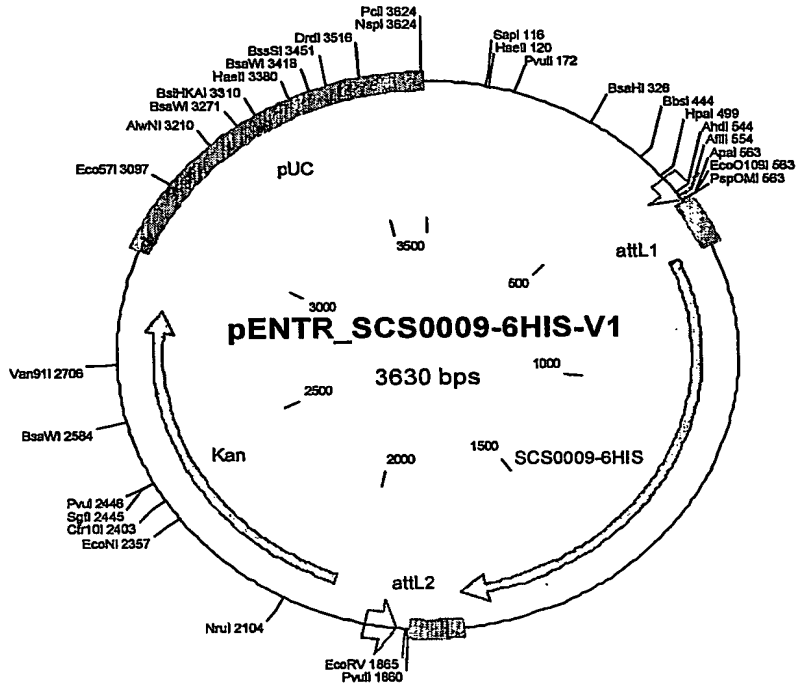


Figure 29

Map of pENTR-SCS0009-6HIS

Molecule:			pENTR_SCS0009-6HIS-V1, 3630 bps DNA Circular		
File Name:			15057[1].cm5,		
Description:			Ligation of SCS0009-6HIS into pdonr221		
Type	Start	End	Name	Description	
MARKER	536			21M13 Forward primer	
REGION	568	670	attL1		
GENE	677	1750	SCS0009-6HIS		
REGION	1754	1855	attL2		
MARKER	1911	C		M13 Reverse primer	
GENE	2024	2833	Kan	Kanamycin resistance gene	
REGION	2954	3627	pUC	pUC origin	



15

Figure 30
Map of pEAK12d-SCS0009-6HIS

Molecule: pEAK12d_SCS0009 -6HIS, 8023 bps DNA Circular
File Name: 15062[1].cm5,
Description: pEAK12 DES with two recombination sites attR1 and attR2 between which the cDNA is inserted

Type	Start	End		Name	Description
REGION	2	595		pmb -ori	
GENE	596	1519		Amp	
REGION	1690	2795		EF -1alpha	
REGION	2796	2845		MCS''	
REGION	2855	2874		attB1	
GENE	2888	3961		S CS0009-6HIS	
REGION	3969	3990		attB2	
REGION	3996	3996		'MCS	
REGION	3997	4425		'A	poly A/splice
GENE	5044	4426	C	PUR	PUROMYCIN
REGION	5268	5045	C	tK	tK promoter
REGION	5763	5269	C	Ori P	
GENE	7815	5763	C	EBNA -1	
REGION	7816	8015		sv40	

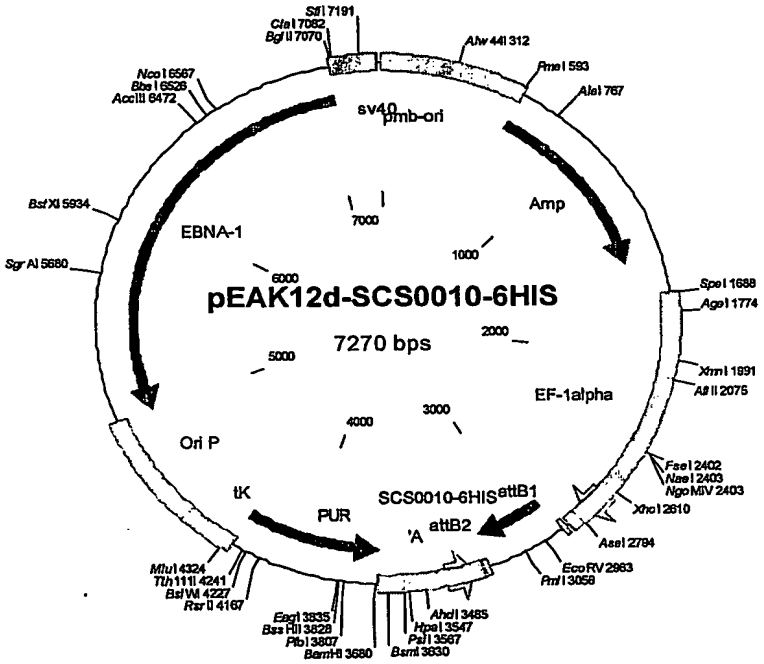


Figure 31

Map of pDEST12.2-SCS0009-6HIS

Molecule: pDEST12.2_SCS0009 -6HIS-V1, 6709 bps DNA Circular
File Name: 15064[1].cm5
Description: Ligation of SCS0009 -6HIS-V1 into pDEST 12.2

Type	Start	End		Name	Description
REGION	15	608		CMV promoter	
REGION	648	665		M13R	Forward primer
REGION	687	706		SP6 promoter	
REGION	730	756		att B1	
GENE	763	1836		SCS0009 -6HIS	
REGION	1865	1840	C	attB2	
GENE	1915	1895	C	T7	T7 promoter
REGION	1943	1926	C	21M13	reverse primer
REGION	2215	2481		pA	SV40 polyadenylation signa l
REGION	2607	3062		f1	f1 intergenic region
REGION	3222	3530		P SV40	SV40 ori & early promoter
GENE	3589	4383		Neo	
REGION	4447	4495		pA	synthetic poly adenylation signal
GENE	4906	5766		Amp	
REGION	5911	6584		ori	pUC ori

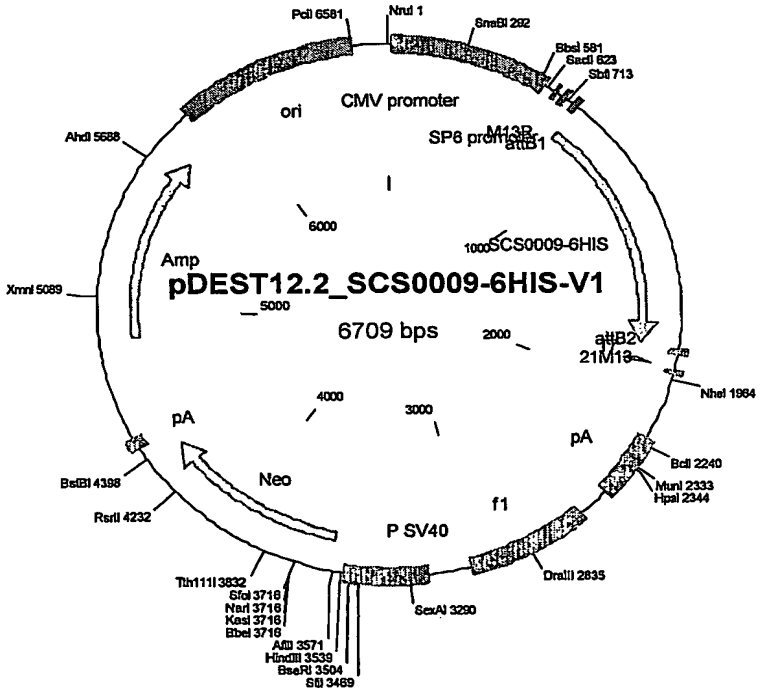
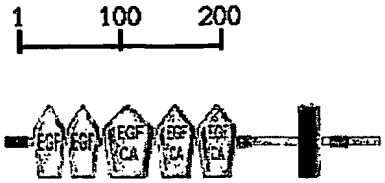
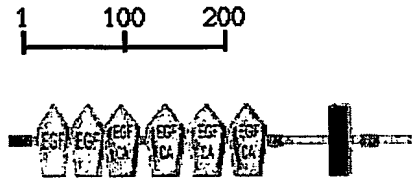


Figure 32

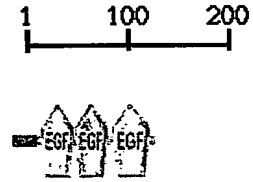
Domains within the query sequence SCS0009 of 352 residues



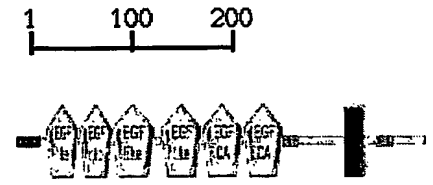
Domains within the query sequence SCS0009-SV3 of 377 residues



Domains within the query sequence SCS0009-SV4 of 134 residues



Domains within the query sequence SCS0009-SV5 of 383 residues is identical to sptrembInew|AAQ88493|AAQ88493



Transmembrane segments as predicted by the *TMHMM2* program (▬), coiled coil regions determined by the *Coils2* program (▬) and Segments of low compositional complexity, determined by the *SEG* program (▬), signal peptides determined by the *Sigcleave* program (▬), GPI anchors are indicated by (▬). Regions containing repeats detected by Prospero, but not covered by domains are indicated by (▬).